

Indici semantici: un catalogo, un solo linguaggio

In questo documento è delineata una proposta per il rinnovamento della catalogazione semantica nella Biblioteca di scienze sociali, e più in generale, nelle biblioteche dell'Università di Firenze. Il fine ultimo di questa proposta è l'istituzione di una pratica coerente ed efficace dell'indicizzazione per soggetto.

La situazione attuale

1. L'indicizzazione per soggetto

L'indicizzazione semantica nella Biblioteca di scienze sociali prevede l'assegnazione ai documenti delle notazioni della *Classificazione decimale Dewey* e delle stringhe di soggetto secondo il *Soggettario* di Firenze.

In realtà, solo la classificazione viene assegnata a tutti i documenti. L'indicizzazione per soggetto è stata abbandonata per molti anni (dalla metà degli anni '90 a circa due anni fa), con l'eccezione dei volumi di filosofia del diritto; attualmente viene effettuata, ma senza sistematicità. Si può pertanto affermare che l'utente della Biblioteca di scienze sociali non può contare su un catalogo per soggetti attendibile.

Al di là dei problemi particolari della Biblioteca di scienze sociali, la stessa pratica dell'indicizzazione per soggetto come viene generalmente condotta nelle biblioteche italiane ha bisogno di un completo rinnovamento. Infatti, se da una parte si riconosce in genere l'opportunità o la necessità di una indicizzazione verbale preordinata, dall'altra è evidente che le tecniche e gli strumenti a disposizione (la soggettazione tradizionale, il cui metodo è sostanzialmente una prassi fondata su criteri di tipo analogico sulla base del *Soggettario*, dei suoi aggiornamenti e della pratica della BNI) non sono adeguati alle necessità dell'organizzazione e ricerca per soggetto dei documenti nel catalogo elettronico. L'arretratezza della soggettazione tradizionale rende difficile per il catalogatore mantenere comportamenti coerenti; inoltre, non sempre le voci di soggetto assegnate ai documenti hanno il requisito della specificità, caratteristica fondamentale di questi indici. La specificità degli indici per soggetto ha un'importanza particolare in una biblioteca universitaria, quindi ad alta specializzazione.

Il percorso dell'atteso rinnovamento dell'indicizzazione per soggetto in Italia appare comunque già tracciato. La *Guida all'indicizzazione per soggetto* del GRIS¹ ha delineato un metodo di indicizzazione non più basato sull'analogia, ma su principi e criteri logici per la sintassi delle voci di soggetto e per la strutturazione del vocabolario controllato. La Biblioteca nazionale centrale di Firenze ha recepito i principi del GRIS nel suo progetto per un *Nuovo soggettario*,² elaborato da un gruppo di lavoro composto in gran parte da esponenti del GRIS.

2. La classificazione

Nella Biblioteca di scienze sociali l'assegnazione del numero di classificazione acquista particolare importanza, perché su di essa è basata la collocazione dei libri a scaffale aperto, che si progetta di estendere a circa la metà del patrimonio librario.

Ogni notazione della CDD esprime, oltre alla classe di appartenenza, uno o più soggetti. Poiché la costruzione del numero CDD è la parte finale di una operazione di indicizzazione che

¹ *Guida all'indicizzazione per soggetto* / Associazione italiana biblioteche. GRIS, Gruppo di ricerca sull'indicizzazione per soggetto. – Roma : Associazione italiana biblioteche, 1996. – Ristampa con correzioni: 2001.

² *Per un nuovo Soggettario : studio di fattibilità sul rinnovamento del Soggettario per i cataloghi delle biblioteche italiane* / Biblioteca nazionale centrale di Firenze ; commissionato dalla BNCF alla IFNET, Firenze ; realizzato dal Gruppo di progetto per il rinnovamento del Soggettario. – Milano : Editrice Bibliografica, 2002.

condivide con la soggettazione le prime fasi (analisi concettuale del documento e creazione della frase titolo), il soggetto o i soggetti contenuti in una particolare notazione hanno un contenuto concettuale in parte simile a quello delle stringhe di soggetto assegnate ai documenti che condividono quel numero di classificazione. Con importanti differenze, che consigliano la creazione dei due diversi tipi di indici semantici, per soggetto classificato e per soggetto verbale:

- (1) i soggetti contenuti nella notazione CDD corrispondono alla notazione, non al documento indicizzato, perché la classificazione spesso non offre all'indicizzatore tutti i concetti necessari per raggiungere il necessario livello di specificità;
- (2) i concetti espressi dalla notazione non sempre coincidono con quelli realmente pertinenti alla descrizione per soggetto di un particolare documento.

In generale, la classificazione ha una duplice funzione: (1) la collocazione dei volumi nello scaffale aperto; (2) l'organizzazione e il recupero dei documenti sulla base del loro contenuto semantico.

Nell'opac di Ateneo gli utenti possono effettuare ricerche sui numeri di classificazione e sugli equivalenti verbali dei numeri di classificazione: insieme alla segnaletica disposta sugli scaffali, l'opac è lo strumento offerto agli utenti per interpretare le notazioni numeriche della CDD, per cui occorre valutarne l'efficacia in relazione alle funzioni della classificazione sopra ricordate.

- (1) Dal punto di vista della funzione di collocazione, la ricerca sulle notazioni numeriche è un utile complemento alla segnaletica sugli scaffali: mentre la segnaletica non può essere particolarmente specifica, una ricerca in opac su una particolare notazione (per esempio una divisione o una sezione della CDD) permette di evidenziare la lista di numeri sempre più lunghi e specifici che si articolano sotto di essa, con gli equivalenti verbali che esplicitano il significato di quelle notazioni.
- (2) Dal punto di vista del recupero semantico, invece, gli attuali equivalenti verbali hanno a mio avviso una utilità assai limitata, perché per essi non è previsto un controllo terminologico: non è verificato che tra concetti e espressioni verbali ci sia una corrispondenza tale da garantire uniformità e univocità, e inoltre è assente la struttura delle relazioni tra concetti (thesaurus) necessaria per permettere che una qualsiasi formulazione dell'utente possa condurlo non solo alla forma preferita per un determinato concetto (e da essa alle notazioni corrispondenti ed alle registrazioni bibliografiche relative), ma anche ai concetti correlati per gerarchia o altrimenti associabili. Dunque la domanda: "sotto quali notazioni si trovano i documenti che trattano un particolare argomento?" non può trovare una risposta puntuale ed esaustiva.

La proposta

Propongo di esprimere il significato delle notazioni della CDD con linguaggio GRIS. In sostituzione degli attuali equivalenti verbali, compilati con termini non controllati, e non coerenti dal punto di vista sintattico perché rispecchiano l'ordine di citazione attribuito convenzionalmente dalla classificazione, ad ogni notazione dovrebbe corrispondere una (o più di una) stringa di soggetto costruita secondo la *Guida* del GRIS, quindi composta da termini di indicizzazione controllati e con una sintassi coerente basata su principi logici. Il controllo dei termini comporta la costruzione di un thesaurus che espliciti tre tipi di relazioni: relazione di equivalenza, relazioni gerarchiche, relazione associativa.

L'indicizzazione per soggetto delle notazioni CDD secondo il metodo GRIS risponde alle seguenti finalità:

- (1) Per l'opac di Ateneo, la possibilità per l'utente di interrogare efficacemente, cioè con esaustività e precisione il catalogo nella sua organizzazione per classi.

- (2) Per la Biblioteca di scienze sociali, la disponibilità, per il patrimonio non soggetto, di stringhe di soggetto coesiste al numero di classificazione assegnato: quindi non sempre con un adeguato livello di specificità, ma comunque sufficienti per permettere agli utenti una ricerca per soggetto sulla totalità del posseduto. Per le ragioni già ricordate, l'indicizzazione per soggetto della CDD può essere utilizzata, senza riprendere i libri in mano, per riparare lo strappo provocato nel catalogo da alcuni anni senza soggetti, ma in generale non può sostituire l'indicizzazione per soggetto dei documenti.
- (3) Per la Biblioteca di scienze sociali e per tutto il Sistema bibliotecario di Ateneo, l'istituzione di *un nuovo linguaggio di indicizzazione*, da utilizzare e arricchire in seguito per l'indicizzazione per soggetto dei documenti al posto della soggettazione tradizionale.

L'istituzione di *un unico linguaggio* per l'indicizzazione per soggetto e per la traduzione verbale dei simboli di classificazione è a mio avviso un passo avanti significativo verso l'effettiva realizzazione, sia dal punto di vista teorico che pratico, dell'unicità del catalogo elettronico, al di là delle diverse strutture organizzative e di reperimento (per autori e titoli, per soggetti, per classi). La reale unicità del catalogo dipende infatti dall'unicità e coerenza del linguaggio catalografico: un linguaggio che dovrebbe essere fondato non su una sommatoria più o meno ragionevole di normative, ma su un impianto coerente, e per questo con una sola semantica, una sola sintassi, una sola pragmatica. Da questo punto di vista molto resta da fare, sia per quanto riguarda il rapporto tra indici per autori e titoli e indici semantici, sia per l'integrazione tra soggetti alfabetici e soggetti classificati. Questa proposta affronta quest'ultimo aspetto, e lo risolve nella creazione di un thesaurus generale per i concetti utilizzati sia nell'indicizzazione per classi, sia nell'indicizzazione alfabetica per soggetto.

Modalità di realizzazione

La Biblioteca di scienze sociali può intraprendere l'indicizzazione per soggetto della CDD limitatamente alle notazioni da essa assegnate ai documenti: attività relativamente semplice dal punto di vista della sintassi, più complessa riguardo alla costruzione del thesaurus, che dell'indicizzazione è parte integrante. Si può prevedere invece che per la completa realizzazione e per il mantenimento di un thesaurus generale per l'indicizzazione semantica come quello qui ipotizzato, non siano sufficienti le forze di una singola biblioteca universitaria, e forse neanche quelle di un sistema di ateneo. Due considerazioni invitano però ad allargare lo sguardo oltre lo stesso Sistema bibliotecario di Ateneo:

- (1) un'indicizzazione per soggetto della CDD con termini controllati sarebbe utile a tutte le biblioteche italiane;
- (2) la costruzione di un thesaurus generale a tre relazioni è prevista nel progetto per il *Nuovo soggettario* della Biblioteca nazionale di Firenze.

Poiché la presente proposta risponde ad esigenze generali, e si inserisce nel processo di rinnovamento dell'indicizzazione in Italia, è realistico perseguire la sua realizzazione mediante l'attivazione di collaborazioni e sinergie.

Nel suo intervento al Congresso AIB (30 ottobre 2003), Anna Lucarelli, in rappresentanza della BNCF, ha dichiarato la disponibilità a collaborare con altre istituzioni, ed ha invitato a presentare idee e progetti.

L'Università di Firenze potrebbe proporre una collaborazione proprio per la strutturazione di un thesaurus per l'indicizzazione per soggetto della CDD, come primo passo verso il thesaurus generale degli indici per soggetto.

Al di là delle considerazioni sulla scarsità delle risorse a disposizione, a mio avviso non è opportuno che un ateneo sviluppi autonomamente un vocabolario controllato, nello stesso periodo

in cui la Nazionale di Firenze si appresta a realizzarne uno. Può invece contribuire alla sua realizzazione, a proprio e generale vantaggio.

Il vocabolario di una lingua è lo stesso per tutti i parlanti, ma ognuno utilizza un proprio sottoinsieme di vocaboli, che può essere assai vasto o ristretto, di uso più o meno comune. Allo stesso modo, un unico vocabolario controllato potrebbe ospitare tutte le espressioni del linguaggio di indicizzazione, anche le più specialistiche, ed essere utilizzato dalle singole biblioteche secondo la necessità. Le biblioteche ad alta specializzazione, come quelle universitarie, hanno necessità di una terminologia più ricca e specifica rispetto alle biblioteche generali, ma questo non significa che la struttura generale del vocabolario debba essere diversa.

La stessa Nazionale di Firenze avrebbe a mio avviso grande vantaggio dal contributo, sotto il suo coordinamento, di diverse istituzioni documentarie ad alta specializzazione all'impresa del *Nuovo soggettario*, perché la produzione editoriale italiana non potrebbe sostenere la costruzione di un vocabolario di indicizzazione sufficientemente ricco: un'incompletezza che lo stesso *Soggettario* lamentava nella sua prefazione, dopo aver sottolineato la rilevanza, per la ricchezza della terminologia, dell'apporto dei libri stranieri: "ben sentiamo che questo soggettario resta molto lontano dal presentare adeguatamente tutti i campi dello scibile" (p. IX).

Il thesaurus generale prescinde dall'adozione di un particolare *software*, e non sarebbe rilevante per la sua creazione e utilizzazione l'adesione presente o futura a SBN. E' necessario invece che il programma che verrà adottato dall'Università di Firenze per la catalogazione sia in grado di gestire un thesaurus a tre relazioni, e nelle stringhe di soggetto di distinguere, quando non siano coincidenti, le rappresentazioni lessicali di un concetto, cioè i termini di indicizzazione, dalle espressioni verbali che in SBN sono dette descrittori.

Conclusioni

Il processo di cambiamento dell'indicizzazione per soggetto in Italia è in corso, e in un futuro ormai prossimo l'attuale *Soggettario* cesserà di essere lo strumento di base per la BNI e per tutte le biblioteche italiane. Si propone di non aspettare "dall'alto" un nuovo metodo al quale ci si dovrà comunque adeguare in futuro, ma di partecipare attivamente alla sua realizzazione.

La Biblioteca di scienze sociali potrebbe avviare l'indicizzazione per soggetto delle notazioni della CDD con una sperimentazione su un campione significativo, per stabilire i criteri da adottare e valutare l'impegno necessario. Un passo ulteriore potrebbe essere il completamento dell'indicizzazione per soggetto delle notazioni attribuite ai documenti della Biblioteca di scienze sociali, oppure delle notazioni della sola classe 300. In seguito, si dovrebbe completare l'indicizzazione delle notazioni attribuite dal Sistema bibliotecario di Ateneo, ed avviare l'indicizzazione per soggetto dei documenti.

In ogni fase successiva alla sperimentazione, la cooperazione, anche interistituzionale, è necessaria.

Firenze, 4 febbraio 2004

Andrea Fabbrizzi